TÉCNICO LISBOA

No Routes Wasted for Waste Collection: Exploring real-time information through ML techniques to improve waste collection

__Tânia Ramos__ [1], Manuel Lopes [2], Carolina Morais [1], Ana Barbosa-Póvoa [1]

[1] CEG-IST, Instituto Superior Técnico, University of Lisbon, Portugal

[3] INESC-ID, Instituto Superior Técnico, University of Lisbon, Portugal

WSMARTROUTE

OPERATION RESEARCH MEETS MACHINE LEARNING
how to get the most of both worlds to achieve excellent Decision Support System

EWG POR PRACTICE OF OR

Introduction

University

TÉCNICO LISBOA

Research Centres :

CEG-IST -> Operations & Logistics; Routing; Optimization

INESC-ID -> Machine Learning

- How to treat the data available?
- How to improve the collection operation based on that data?

Technological Company

evox Inspiring Innovation

Ultrasonic sensors

360 WASTE

Dashboard

18 sensorized bins

Municipality

CÂMARA MUNICIPAL CASTELO BRANCO

Objective

- Define smart collection routes that maximize operational profit while avoiding bins' overflow using real-time data on the bins' fill-level (transmitted by ultrasonic sensors placed inside the bins)

  - Adequate treatment of real time data
  - Development of optimization models that account for these real time data

How to treat the data available?

EWG POR
PRACTICE OF OR

- 18 sensorized bins since September 2017;
- Data is transmitted whenever a significative variation occurs at the bins' fill-level: Several transmissions per day.

Descrição	Actual Volume	%/day	Evolution	Last Reading
48843 Contentor 607 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	68%	39		19/06/2018 20:35
15415 Contentor 611 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	100%	39		19/06/2018 18:38
49619 Contentor 603 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	48%	29		19/06/2018 15:16
52910 Contentor 606 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	31%	16		19/06/2018 11:06
50419 Subterrâneo 3 - 616 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	51%	23		19/06/2018 08:29

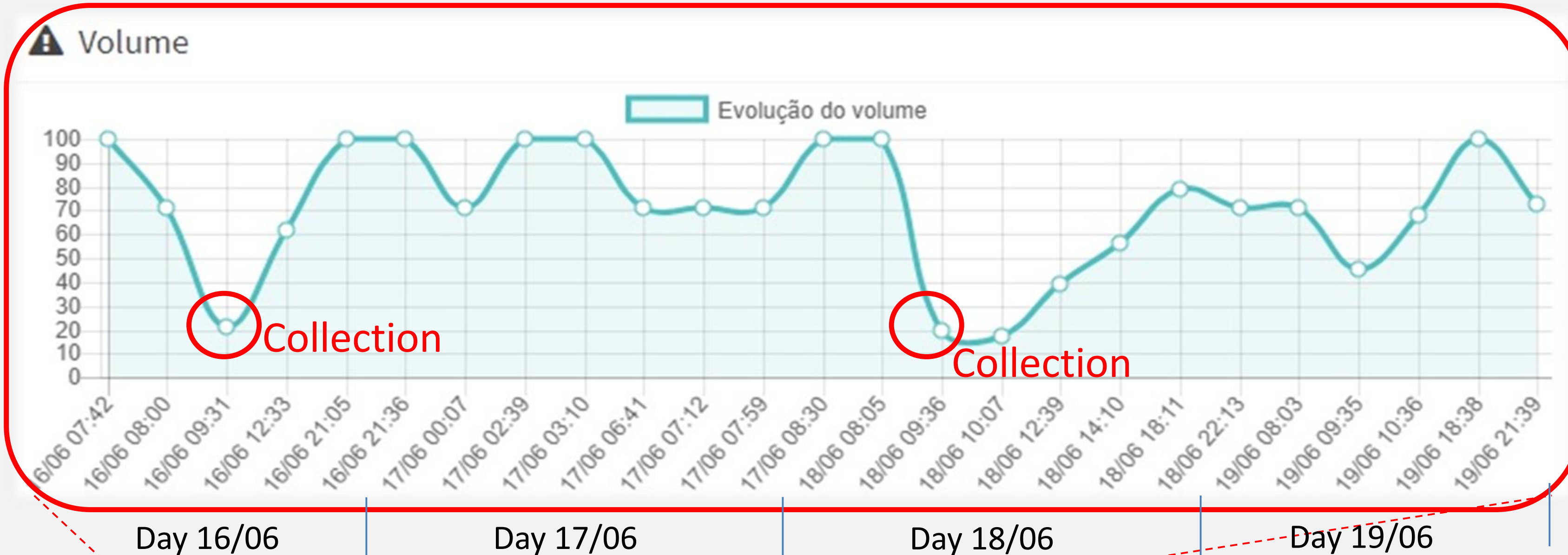
Source: Extracted from 360Waste Platform (www.360waste.pt)

How to treat the data available?

Day 16/06 Day 17/06 Day 18/06 Day 19/06

Descrição	Volume Atual	%/dia	Evolução	Última Leitura
48843 Contentor 607 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	68%	39		19/06/2018 20:35
15415 Contentor 611 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	100%	39		19/06/2018 18:38
49619 Contentor 603 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	48%	29		19/06/2018 15:16
52910 Contentor 606 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	31%	16		19/06/2018 11:06
50419 Subterrâneo 3 - 616 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	51%	23		19/06/2018 08:29

How to treat the data available?



Descrição	Volume Atual	%/dia	Evolução	Última Leitura
48843 Contentor 607 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	68%	39		19/06/2018 20:35
15415 Contentor 611 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	100%	39		19/06/2018 18:38
49619 Contentor 603 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	48%	29		19/06/2018 15:16
52910 Contentor 606 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	31%	16		19/06/2018 11:06
50419 Subterrâneo 3 - 616 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	51%	23		19/06/2018 08:29

How to treat the data available?

Day 16/06 | Day 17/06 | Day 18/06 | Day 19/06

Descrição	Volume Atual	%/dia	Evolução	Última Leitura
48843 Contentor 607 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	68%	39		19/06/2018 20:35
15415 Contentor 611 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	100%	39		19/06/2018 18:38
49619 Contentor 603 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	48%	29		19/06/2018 15:16
52910 Contentor 606 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	31%	16		19/06/2018 11:06
50419 Subterrâneo 3 - 616 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	51%	23		19/06/2018 08:29

How to treat the data available?

Day 16/06 Day 17/06 Day 18/06 Day 19/06

Descrição	Volume Atual	%/dia	Evolução	Última Leitura
48843 Contentor 607 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	68%	39		19/06/2018 20:35
15415 Contentor 611 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	100%	39		19/06/2018 18:38
49619 Contentor 603 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	48%	29		19/06/2018 15:16
52910 Contentor 606 - Exterior 800 litros - Polietileno - Resíduos Sólidos Urbanos	31%	16		19/06/2018 11:06
50419 Subterrâneo 3 - 616 - Exterior 1000 litros - Polietileno - Resíduos Sólidos Urbanos	51%	23		19/06/2018 08:29

IMPORTANT: adequate treatment of the available data.

How to treat the data available?

EWG POR
PRACTICE OF OR

Machine Learning

Sensors:

- Sensors give a "noisy" measure of the volume as the height inside of the container is not uniform;
- Volume can reduce when heavier objects are inserted.

Container:

- We measure volume but the total volume changes stochastically with time; and **we want to avoid containers' overflowing.**

Predict when a container will be completely full based on noisy sensors

How?

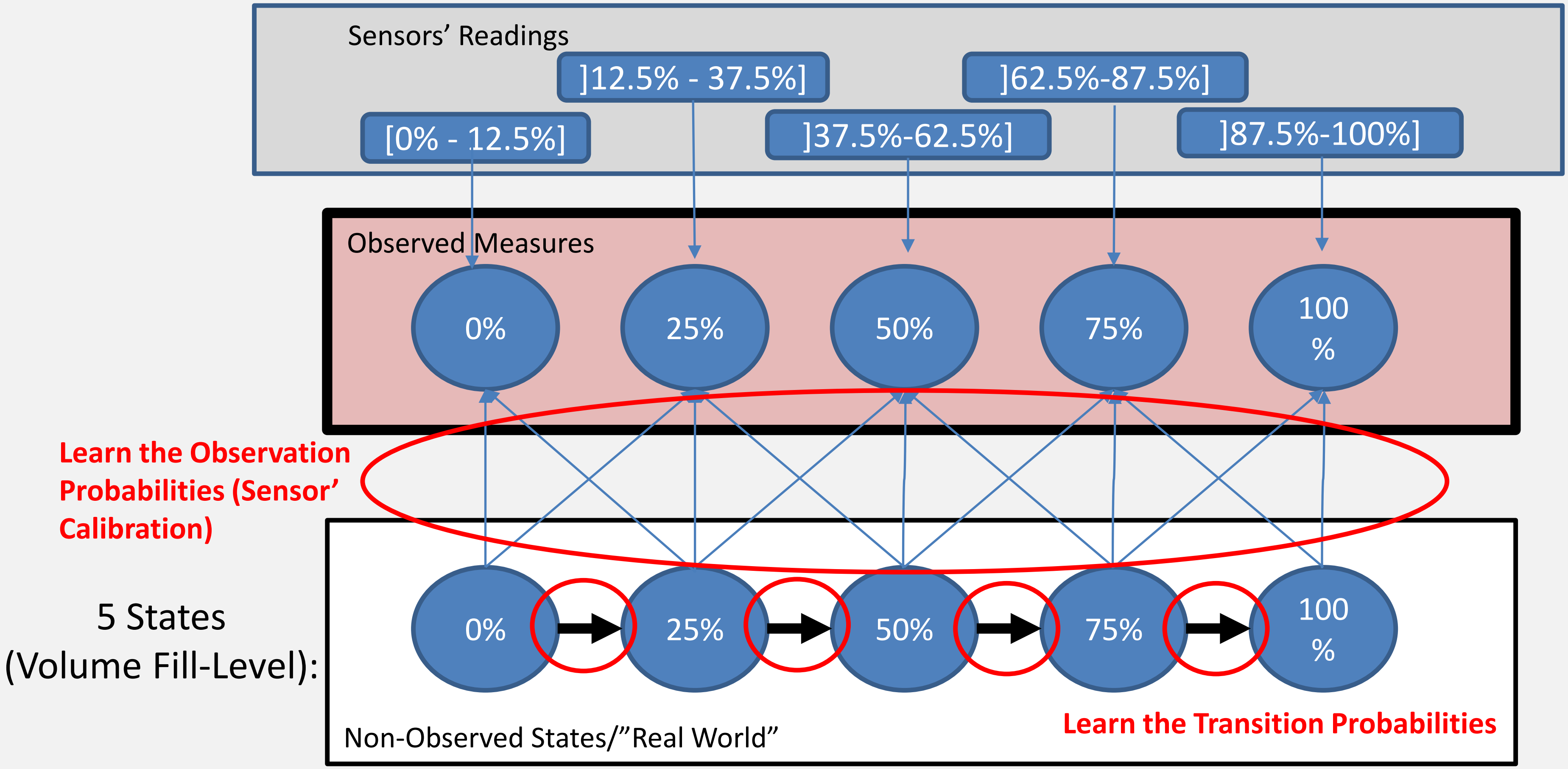
Hidden Markov Models (HMM)

Probabilities of becoming full are learnt from the observations

How to treat the data available?

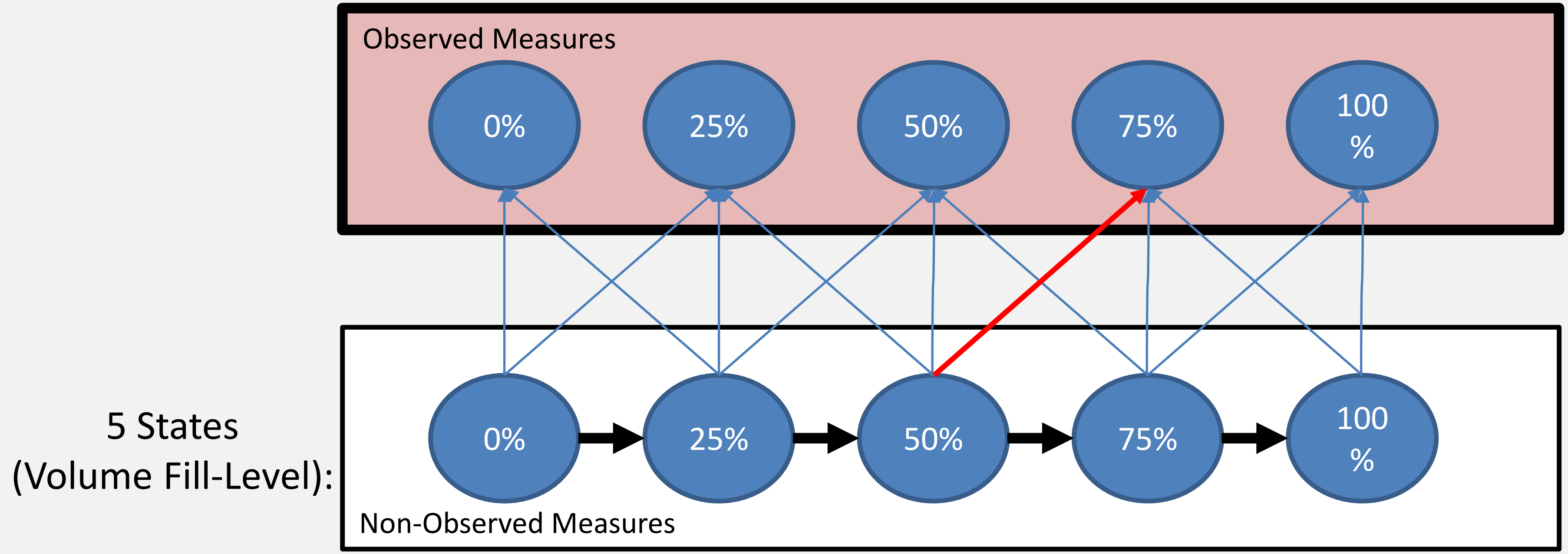
EWG POR
PRACTICE OF OR

Hidden Markov Models (HMM)



How to treat the data available?

Hidden Markov Models (HMM)



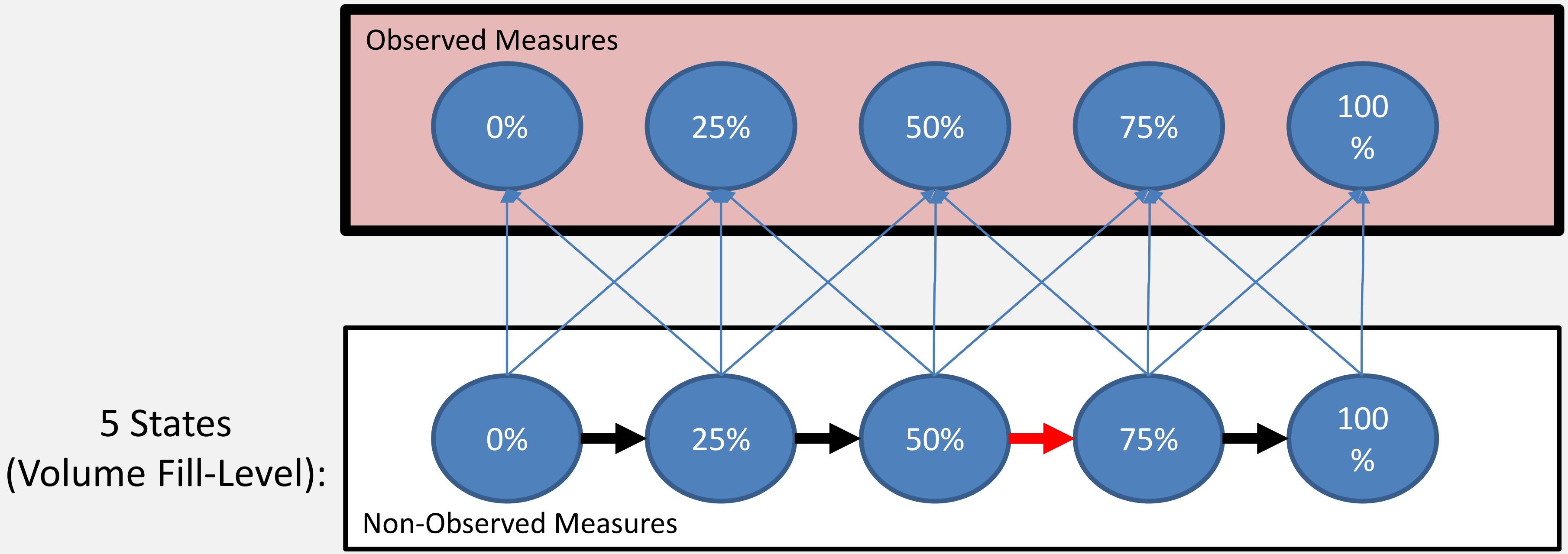
Probability of Observation

P(Observation|State)

e.g. P(Observation = 75% | State =50%) =0.1 -> Probability of 10% of measuring 75% filling when in reality it is just 50%

How to treat the data available?

Hidden Markov Models (HMM)



Probability of Transition

$P(State_{t+1}|State_t)$

e.g. $P(State_{t+1}=75\%|State_t=50\%) = 0.8$ -> Probability of 80% to change from state 50% to stage 75% in the next day

How to treat the data available?

EWG POR PRACTICE OF OR

Hidden Markov Models (HMM)

e.g.

Bin 601

Probability of each state at each day in the future, given the observation measure $p(x_{t+d}|o_t)$

Observed Measure: 0%

	States				
DAYS	**0%**	**25%**	**50%**	**75%**	**100%**
0	0.875	0.125	0	0	0
1	0.08643	0.237477	0.296044	0.22099	0.159059
2	0.008537	0.045695	0.116409	0.187722	0.641637
3	0.000843	0.00671	0.02583	0.064063	0.902553
4	8.33E-05	0.00088	0.004532	0.015174	0.97933
5	8.23E-06	0.000108	0.000699	0.002948	0.996237
6	8.13E-07	1.28E-05	9.94E-05	0.000505	0.999381
7	8.03E-08	1.48E-06	1.34E-05	7.95E-05	0.999906

Observed Measure: 25%

	States				
DAYS	**0%**	**25%**	**50%**	**75%**	**100%**
0	0.2	0.7	0.1	0	0
1	0.019756	0.120603	0.250297	0.27873	0.330614
2	0.001951	0.016996	0.061712	0.132441	0.7869
3	0.000193	0.002181	0.011057	0.034494	0.952075
4	1.90E-05	0.000265	0.001711	0.00695	0.991054
5	1.88E-06	3.11E-05	0.000243	0.001211	0.998513
6	1.86E-07	3.55E-06	3.26E-05	0.000192	0.999772
7	1.84E-08	3.99E-07	4.19E-06	2.84E-05	0.999967

Observed Measure: 50%

	States				
DAYS	**0%**	**25%**	**50%**	**75%**	**100%**
0	0	0.111111	0.777778	0.111111	0
1	0	0.010975	0.105415	0.244599	0.63901
2	0	0.001084	0.013236	0.054593	0.931086
3	0	0.000107	0.001586	0.009125	0.989181
4	0	1.06E-05	0.000184	0.001342	0.998463
5	0	1.04E-06	2.09E-05	0.000183	0.999795
6	0	1.03E-07	2.34E-06	2.38E-05	0.999974
7	0	1.02E-08	2.57E-07	2.98E-06	0.999997

Observed Measure:75%

	States				
DAYS	**0%**	**25%**	**50%**	**75%**	**100%**
0	0	0	0.111111	0.777778	0.111111
1	0	0	0.010975	0.105415	0.88361
2	0	0	0.001084	0.013236	0.985679
3	0	0	0.000107	0.001586	0.998307
4	0	0	1.06E-05	0.000184	0.999805
5	0	0	1.04E-06	2.09E-05	0.999978
6	0	0	1.03E-07	2.34E-06	0.999998
7	0	0	1.02E-08	2.57E-07	1

Observed Measure: 100%

	States				
DAYS	**0%**	**25%**	**50%**	**75%**	**100%**
0	0	0	0	0.125	0.875
1	0	0	0	0.012347	0.987653
2	0	0	0	0.00122	0.99878
3	0	0	0	0.00012	0.99988
4	0	0	0	1.19E-05	0.999988
5	0	0	0	1.18E-06	0.999999
6	0	0	0	1.16E-07	1
7	0	0	0	1.15E-08	1

How to treat the data available?

EWG POR
PRACTICE OF OR

Hidden Markov Models (HMM)

e.g.

Bin 601

Probability of each state at each day in the future, given the observation measure $p(x_{t+d}|o_t)$

Observed Measure: 25%

	States				
DAY	**0%**	**25%**	**50%**	**75%**	**100%**
0	0.2	0.7	0.1	0	0
1	0.019756	0.120603	0.250297	0.27873	0.330614
2	0.001951	0.016996	0.061712	0.132441	0.7869
3	0.000193	0.002181	0.011057	0.034494	0.952075
4	1.90E-05	0.000265	0.001711	0.00695	0.991054
5	1.88E-06	3.11E-05	0.000243	0.001211	0.998513
6	1.86E-07	3.55E-06	3.26E-05	0.000192	0.999772
7	1.84E-08	3.99E-07	4.19E-06	2.84E-05	0.999967

If the observed measure is 25%, there is a probability of 70% of the actual state to be 25%, 10% of the actual state to be 50% and 20% of the actual state to be 0%

How to treat the data available?

EWG POR
PRACTICE OF OR

Hidden Markov Models (HMM)

e.g.

Bin 601

Probability of the state at each day in the future given the observation measure $p(x_{t+d}|o_t)$

Observed Measure: 25%

	States				
DAY	**0%**	**25%**	**50%**	**75%**	**100%**
0	0.2	0.7	0.1	0	0
1	0.019756	0.120603	0.250297	0.27873	0.330614
2	0.001951	0.016996	0.061712	0.132441	0.7869
3	0.000193	0.002181	0.011057	0.034494	0.952075
4	1.90E-05	0.000265	0.001711	0.00695	0.991054
5	1.88E-06	3.11E-05	0.000243	0.001211	0.998513
6	1.86E-07	3.55E-06	3.26E-05	0.000192	0.999772
7	1.84E-08	3.99E-07	4.19E-06	2.84E-05	0.999967

If the observed measure is 25%, there is a probability of 12% of staying in that state in the next day, 25% of changing to the state 50%, 28% of changing to state 75% and 33% of becoming full in the next day

How to treat the data available?

EWG POR
PRACTICE OF OR

Hidden Markov Models (HMM)

e.g.

Bin 601

	Initial Morning Stock
2/Apr	62%

Sensors' information

Observed Measure
50%

If the bin is 50% full today, in how many days it will be completely full (100%)?

Probability Matrix from HMM

	Observed Measure: 50%				
	States				
DAYS	0%	25%	50%	75%	100%
0	0	0.111111	0.777778	0.111111	0
1	0	0.010975	0.105415	0.244599	0.63901
2	0	0.001084	0.013236	0.054593	0.931086
3	0	0.000107	0.001586	0.009125	0.989181
4	0	1.06E-05	0.000184	0.001342	0.998463
5	0	1.04E-06	2.09E-05	0.000183	0.999795
6	0	1.03E-07	2.34E-06	2.38E-05	0.999974
7	0	1.02E-08	2.57E-07	2.98E-06	0.999997

Bin 601 will be completely full (100%) after 2 days, with 93.1% of probability

Bin 601 will be completely full (100%) after 3 days, with 98.9% of probability

How to treat the data available?

EWG POR
PRACTICE OF OR

Hidden Markov Models (HMM)

e.g.

Bin 601

	Initial Morning Stock
2/Apr	62%

Sensors' information

Observed Measure
50%

Probability Matrix from HMM

	50%				
	States				
DAYS	0%	25%	50%	75%	100%
0	0	0.111111	0.777778	0.111111	0
1	0	0.010975	0.105415	0.244599	0.63901
2	0	0.001084	0.013236	0.054593	0.931086
3	0	0.000107	0.001586	0.009125	0.989181
4	0	1.06E-05	0.000184	0.001342	0.998463
5	0	1.04E-06	2.09E-05	0.000183	0.999795
6	0	1.03E-07	2.34E-06	2.38E-05	0.999974
7	0	1.02E-08	2.57E-07	2.98E-06	0.999997

Daily Accumulation Rate = (1 – 0.5) / 2 days 25%/day

Daily Accumulation Rate = (1 – 0.5) / 3 days 17%/day

Need to define a Confidence Level Threshold **95%**

Bin 601 will be completely full after 2 days, with 93.1% of probability

Bin 601 will be completely full after 3 days, with 98.9% of probability

How to improve the collection operation based on that data?

EWG POR
PRACTICE OF OR

Smart Waste Collection Routing Problem *(Ramos et al. 2018)*

Use of **real-time information** on the bins' fill-level (transmitted by volumetric sensors placed inside the bins) to define **smart collection routes** that **maximize operational profit:**

Max **PROFIT** = **revenues** obtained from the recyclable waste collected - **transportation costs** of collecting that waste

Maximize the amount of waste collected while minimizing distance travelled

How to improve the collection operation based on that data?

EWG POR
PRACTICE OF OR

Smart Waste Collection Routing Problem *(Ramos et al. 2018)*

Decision: To select the **waste bins to be visited** (if any) and the **optimal visiting sequence** in each ***day t*** for each ***vehicle k***, which will **maximize the profit** while satisfying the vehicles' capacity, the bins' capacity and a service level (measured by the number of overflowing bins).

How to improve the collection operation based on that data?

Decision: To select the **waste bins to be visited** (if any) and the **optimal visiting sequence** in each ***day t*** for each ***vehicle k***, which will **maximize the profit** while satisfying the vehicles' fixed capacity, the bins' capacity and a service level.

VRP with Profit (VRPP)[1] :
Maximize Profit for One Day

VRPP Model is solved every day *t*, in the morning, after receiving sensors' data on the bins' fill-level.

Problem: "Blind" to future events

[1](Ramos et al. 2018)

Inventory Routing Problem (IRP):
Maximize Profit for a Time Horizon

Static IRP

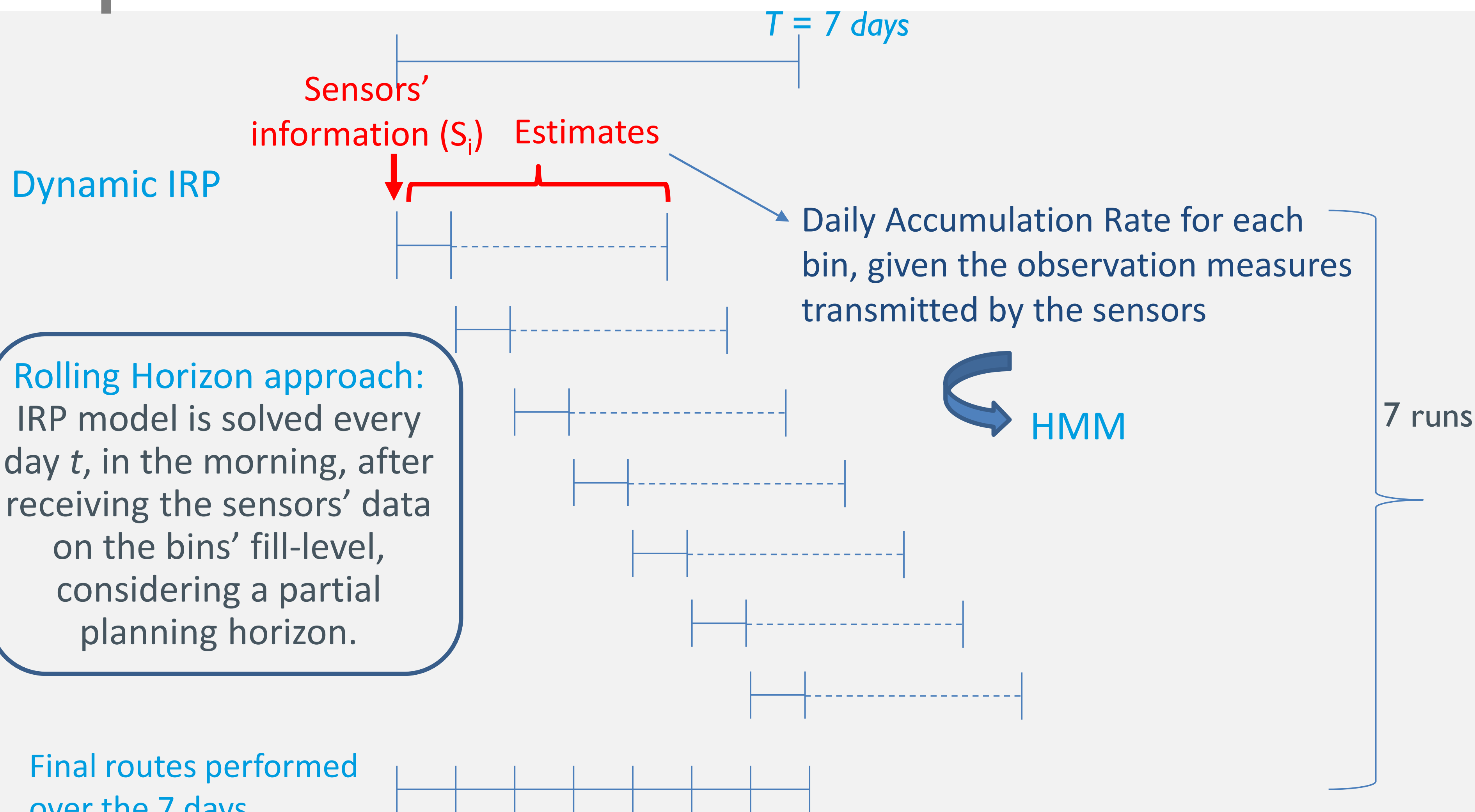
IRP model is solved at day t=1, in the morning, after receiving sensors' data, considering the entire planning horizon (e.g., 7 days).

Problem: Considers real-time data only for the first day (deals with estimates for the days ahead).

Dynamic IRP

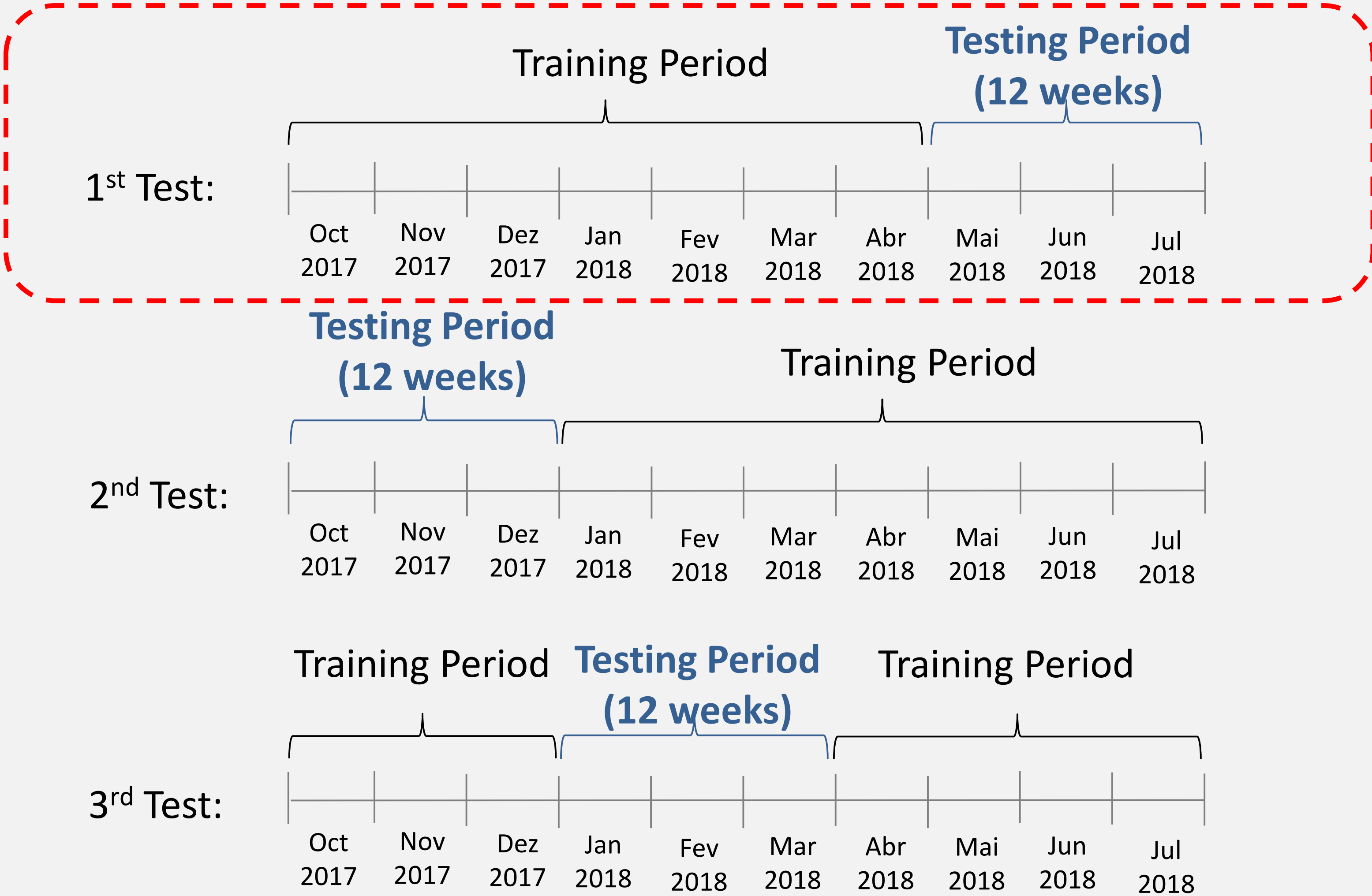
Considers a continuous data updating in the model.

How to improve the collection operation based on that data?



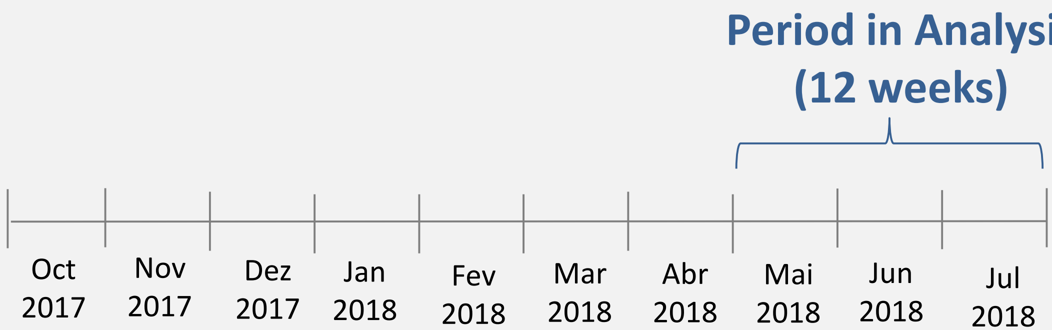
Rolling Horizon approach: IRP model is solved every day *t*, in the morning, after receiving the sensors' data on the bins' fill-level, considering a partial planning horizon.

Testing Set



Training Period
Testing Period (12 weeks)
1st Test:
Oct 2017
Nov 2017
Dez 2017
Jan 2018
Fev 2018
Mar 2018
Abr 2018
Mai 2018
Jun 2018
Jul 2018
Testing Period (12 weeks)
Training Period
2nd Test:
Oct 2017
Nov 2017
Dez 2017
Jan 2018
Fev 2018
Mar 2018
Abr 2018
Mai 2018
Jun 2018
Jul 2018
Training Period
Testing Period (12 weeks)
Training Period
3rd Test:
Oct 2017
Nov 2017
Dez 2017
Jan 2018
Fev 2018
Mar 2018
Abr 2018
Mai 2018
Jun 2018
Jul 2018

Results – Current Situation



- All 18 bins are collected every Monday, Thursday and Saturday, regardless its fill-level

KPI	TOTAL (12 weeks)	AVERAGE (week)
Weight (kg)	34 403	2 867
Distance (km)	3 584	299
Attended bins	648	54
Full bins (87,5%<S_i<100%)	0	0
Overflowing bins (S_i>100%)	235	20
Ratio (kg/km)	9.6	9.6

37% of the waste bins are overflowing

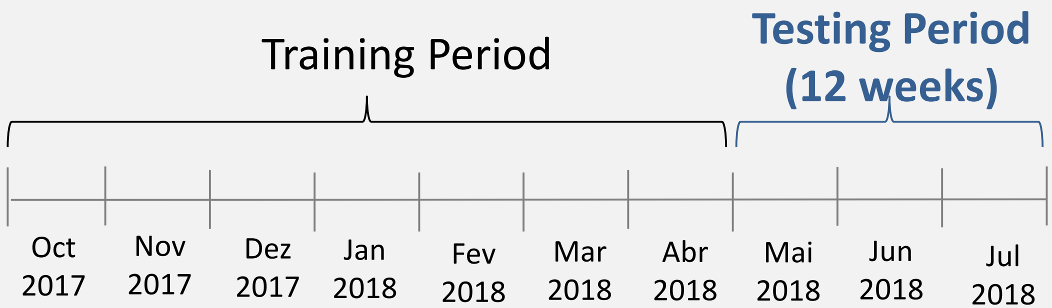
Poor Service Level

High Efficiency

Results – Dynamic IRP w/ HMM

EWG POR
PRACTICE OF OR

1st Test:



- The expected daily accumulation rate computed through the probability matrix from the HMM feeds the Dynamic IRP model

KPI	TOTAL (12 weeks)	AVERAGE (week)
Weight (kg)	38 753	3 229
Distance (km)	3 829	319
Attended bins	468	39
Full bins (87,5%<S_i<100%)	206	17
Overflowing bins ($S_i \geq$100%)	69	6
Ratio (kg/km)	10.1	10.1

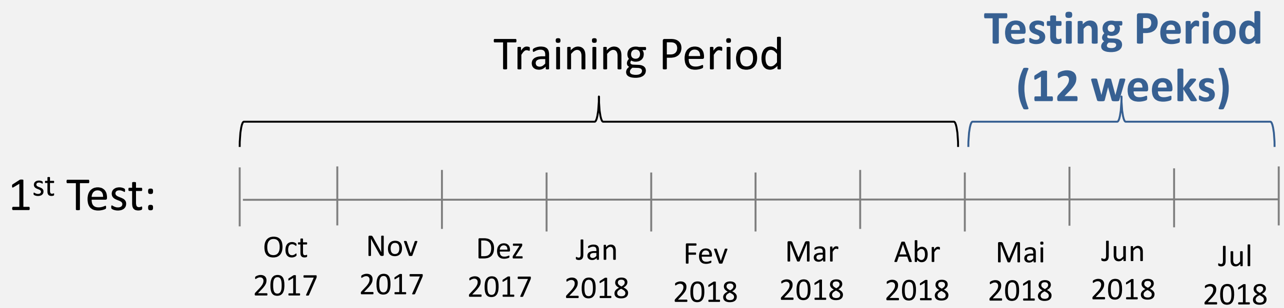
15% of the waste bins are overflowing

Better Service Level

Higher Efficiency

Results – Dynamic IRP w/ HMM

EWG POR
PRACTICE OF OR

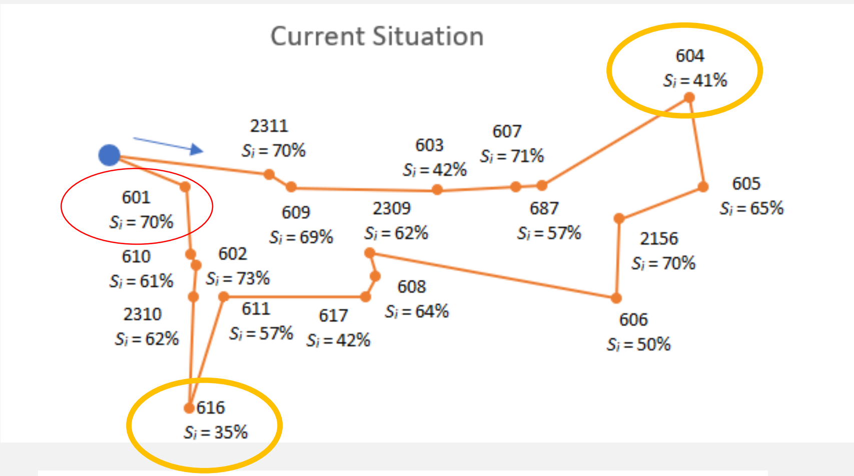


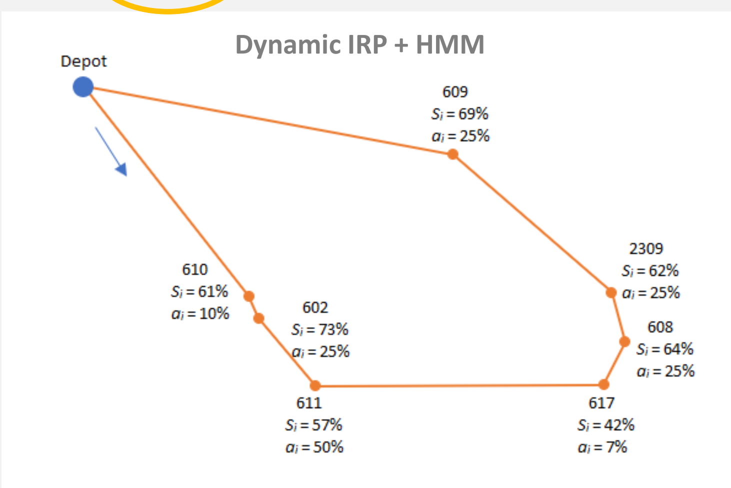
June – 1st week								
KPI	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Total
Weight (kg)	578.54	392.92	0.00	493.74	530.22	241.57	562.76	2799.75
Distance (km)	57.23	47.88	0.00	64.53	57.23	47.31	57.23	331.41
Attended bins	7	5	0	5	6	3	7	33
Full bins (87.5%<Si<100%)	4	2	1	2	3	1	2	15
Overflowing bins (Si>100%)	0	0	0	1	0	0	0	1
Ratio (kg/km)	10.11	8.21	0.00	7.65	9.26	5.11	9.83	8.45

July – 3rd week								
KPI	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Total
Weight (kg)	461.69	1651.55	26.62	11.09	1469.88	150.37	1061.7	4832.9
Distance (km)	37.53	82.93	20.84	8.30	98.79	21.59	63.1	333.1
Attended bins	7	17	1	1	16	3	14	59
Full bins (87.5%<Si<100%)	0	0	0	0	4	0	2	6
Overflowing bins (Si>100%)	0	7	0	0	3	0	0	10
Ratio (kg/km)	12.30	19.92	1.28	1.34	14.88	6.96	16.8	14.5

Results – Routes

e.g. July – 3rd Week, Day 1 (16/July)





e.g.

Bin 601

	Sensors' Reading (8 am)	HMM Daily Accumulation Rate	Actual Daily Accumulation Rate
16/Jul	70%	25%	65%

The IRP model chooses not to collect bin 601 at day 1 (where 70% of volume would be collected); it schedules bin 601 to day 2 (where 95% of volume would be collected)

But… Actual Accumulation Rate for bin 601 on day 16/Jul = 65%

Overflowing bin → 17/Jul = Sensors' Reading = 135% (70%+65%)

Conclusions

KPIs	Current Situation	Dynamic IRP + HMM	
Total weight (kg)	34 403	38 753	+ 13%
Total distance (km)	3 584	3 829	+ 6%
Total attended bins	648	468	- 28%
Total overflowing bins (Si≥100%)	235	69	- 71%
Ratio (kg/km)	9.6	10.1	+ 5%

- **Current Situation -> Efficient (high kg/km ratio), but… 235 overflowing bins!**

  3 routes/week

- **Dynamic IRP + HMM -> Reduces the number of overflowing bins in 71% (69 vs. 235) and increases the efficiency in 5% (10.1 kg/km vs. 9.6 kg/km).**

  5.7 routes/week (average)

Further Work

- **Data Treatment**

  Investing more time on Machine Learning techniques to treat properly the sensors' data

  - Increase the number of states in the HMM?
  - Test different confidence levels
  - Learning the Daily Accumulation Rates (using AutoRegressive models like ARX or ARMA) instead of learning the probabilities of becoming full (HMM)

- **Routing Plan – Dynamic IRP**

  Small instances were tested (18 bins, planning period of 7 days)

  - Develop other solution methods to solve larger instances (matheuristics, metaheuristics, …)

  Improve the integration between HMM and IRP model -> Stochastic IRP?

THANK YOU FOR YOUR ATTENTION!

WSMARTROUTE

http://wsmartroute.tecnico.ulisboa.pt/

This work is part of the WSmart Route Project funded by MIT Portugal/FCT (Fundação para a Ciência e Tecnologia) - MIT-EXPL/SUS/0132/2017

MIT Portugal

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

Tânia Ramos | *tania.p.ramos@tecnico.ulisboa.pt*
Manuel Lopes | *manuel.lopes@tecnico.ulisboa.pt*
Carolina Morais | *carolina.morais@tecnico.ulisboa.pt*
Ana Barbosa-Póvoa | *apovoa@tecnico.ulisboa.pt*